



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 110/708,952

Source: EFS - OLPE -

Date Processed by STIC: 4-5-04

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

**FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221**

**Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:**

**<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>>., EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 4B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03

Does Not Comply  
Corrected Diskette Needed

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER: 10/708,952

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1        Wrapped Nucleics  
    Wrapped Aminos   The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2        Invalid Line Length   The rules require that a line **not exceed** 72 characters in length. This includes white spaces.
- 3        Misaligned Amino  
    Numbering       The numbering under each 5<sup>th</sup> amino acid is misaligned. Do **not** use tab codes between numbers; use **space characters**, instead.
- 4        Non-ASCII       The submitted file was **not** saved in ASCII(DOS) text, as **required** by the Sequence Rules. Please **ensure your subsequent submission is saved in ASCII text**.
- 5        Variable Length   Sequence(s)        contain n's or Xaa's representing more than one residue. **Per Sequence Rules, each n or Xaa can only represent a single residue.** Please present the **maximum** number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6        PatentIn 2.0  
    "bug"           A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)       . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.**
- 7        Skipped Sequences  
    (OLD RULES)   Sequence(s)        missing. If intentional, please insert the following lines for **each** skipped sequence:  
                  (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
                  (i)       SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
                  (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
                  This sequence is intentionally skipped  
  
                  Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to **include** the skipped sequences.
- 8        Skipped Sequences  
    (NEW RULES)   Sequence(s)        missing. If intentional, please insert the following lines for **each** skipped sequence.  
                  <210> sequence id number  
                  <400> sequence id number  
                  000
- 9        Use of n's or Xaa's  
    (NEW RULES)   Use of n's and/or Xaa's have been detected in the Sequence Listing.  
                  Per 1.823 of Sequence Rules, use of <220>-<223> is **MANDATORY** if n's or Xaa's are present.  
                  In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10        Invalid <213>  
    Response       Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is **required** when <213> response is Unknown or is Artificial Sequence
- 11        Use of <220>       Sequence(s)        missing the <220> "Feature" and associated numeric identifiers and responses.  
                  Use of <220> to <223> is **MANDATORY** if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
                  (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12        PatentIn 2.0  
    "bug"           Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13        Misuse of n/Xaa   "n" can **only** represent a single nucleotide; "Xaa" can **only** represent a single amino acid

CRFI



Does Not Comply  
Corrected Diskette Needed

IFWO

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/708,952

DATE: 04/05/2004

TIME: 09:41:15

Input Set : N:\efs\10708952\_efs\SEQ\_LIST.txt

Output Set: N:\CRF4\04052004\J708952.raw

3 <110> APPLICANT: ROSETTA GENOMICS LTD  
 5 <120> TITLE OF INVENTION: BIOINFORMATICALLY DETECTABLE GROUP OF NOVEL REGULATORY VIRAL  
 6 AND VIRAL ASSOCIATED OLIGONUCLEOTIDES AND USES THEREOF  
 8 <130> FILE REFERENCE: 55035  
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/708,952  
 C--> 10 <141> CURRENT FILING DATE: 2004-04-02  
 10 <160> NUMBER OF SEQ ID NOS: 399737  
 12 <170> SOFTWARE: PatentIn version 3.2

## ERRORED SEQUENCES

1002677 <210> SEQ ID NO: 111408  
 1002678 <211> LENGTH: 23  
 1002679 <212> TYPE: RNA  
 1002680 <213> ORGANISM: Human adenovirus E  
 1002682 <400> SEQUENCE: 111408  
 E--> 1002683 ucauaug<sup>g</sup>gcaguuacac cug  
 1002776 <210> SEQ ID NO: 111419  
 1002777 <211> LENGTH: 19  
 1002778 <212> TYPE: RNA  
 1002779 <213> ORGANISM: Human adenovirus E  
 1002781 <400> SEQUENCE: 111419  
 E--> 1002782 uccagcu<sup>u</sup>ccaacccuu  
 1005170 <210> SEQ ID NO: 111685  
 1005171 <211> LENGTH: 17  
 1005172 <212> TYPE: RNA  
 1005173 <213> ORGANISM: Human adenovirus E  
 1005175 <400> SEQUENCE: 111685  
 E--> 1005176 aguugc<sup>n</sup>aga uggccac  
 1005800 <210> SEQ ID NO: 111755  
 1005801 <211> LENGTH: 20  
 1005802 <212> TYPE: RNA  
 1005803 <213> ORGANISM: Human adenovirus E  
 1005805 <400> SEQUENCE: 111755  
 E--> 1005806 gccgaguugc <sup>n</sup>agauggcca  
 1006106 <210> SEQ ID NO: 111789  
 1006107 <211> LENGTH: 21  
 1006108 <212> TYPE: RNA  
 1006109 <213> ORGANISM: Human adenovirus E  
 1006111 <400> SEQUENCE: 111789  
 E--> 1006112 gcuccagcu<sup>n</sup>uuccaaccuu u  
 1006664 <210> SEQ ID NO: 111851

A feature <220> - <223> is mandatory  
 when 'n' is used as a nucleotide  
 designator. Please see rule # 9  
 on error summary sheet.  
 Please check entire Sequence Listing  
 for the above error.

Same

19

Same

17

Same

20

Same

21

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/708,952

DATE: 04/05/2004

TIME: 09:41:50

Input Set : N:\efs\10708952\_efs\SEQ\_LIST.txt

Output Set: N:\CRF4\04052004\J708952.raw

```

1006665 <211> LENGTH: 22
1006666 <212> TYPE: RNA
1006667 <213> ORGANISM: Human adenovirus E
1006669 <400> SEQUENCE: 111851
E--> 1006670 gauggugcan gaguguaaca ug 22
1009283 <210> SEQ ID NO: 112142
1009284 <211> LENGTH: 14
1009285 <212> TYPE: RNA
1009286 <213> ORGANISM: Human adenovirus E
1009288 <400> SEQUENCE: 112142
E--> 1009289 ugcan gagug uaac 14
1010750 <210> SEQ ID NO: 112305
1010751 <211> LENGTH: 18
1010752 <212> TYPE: RNA
1010753 <213> ORGANISM: Human adenovirus E
1010755 <400> SEQUENCE: 112305
E--> 1010756 gaguugan agggccac 18
1012172 <210> SEQ ID NO: 112463
1012173 <211> LENGTH: 15
1012174 <212> TYPE: RNA
1012175 <213> ORGANISM: Human adenovirus E
1012177 <400> SEQUENCE: 112463
E--> 1012178 cgcaan uaca ccccc 15
1019966 <210> SEQ ID NO: 113329
1019967 <211> LENGTH: 19
1019968 <212> TYPE: RNA
1019969 <213> ORGANISM: Human adenovirus E
1019971 <400> SEQUENCE: 113329
E--> 1019972 uccccan uac uaacaacca 19
1020929 <210> SEQ ID NO: 113436
1020930 <211> LENGTH: 22
1020931 <212> TYPE: RNA
1020932 <213> ORGANISM: Human adenovirus E
1020934 <400> SEQUENCE: 113436
E--> 1020935 cugaauuaug an auaaag aa 22
1021937 <210> SEQ ID NO: 113548
1021938 <211> LENGTH: 20
1021939 <212> TYPE: RNA
1021940 <213> ORGANISM: Human adenovirus E
1021942 <400> SEQUENCE: 113548
E--> 1021943 gcggccgan gccugggcga 20
2603003 <210> SEQ ID NO: 289222
2603004 <211> LENGTH: 20
2603005 <212> TYPE: RNA
2603006 <213> ORGANISM: Human adenovirus E
2603008 <400> SEQUENCE: 289222
E--> 2603009 gccgaguugc an agaugcca 20
2605865 <210> SEQ ID NO: 289540
2605866 <211> LENGTH: 17

```

## RAW SEQUENCE LISTING

DATE: 04/05/2004

PATENT APPLICATION: US/10/708,952

TIME: 09:42:55

Input Set : N:\efs\10708952\_efs\SEQ\_LIST.txt

Output Set: N:\CRF4\04052004\J708952.raw

```

2605867 <212> TYPE: RNA
2605868 <213> ORGANISM: Human adenovirus E
2605870 <400> SEQUENCE: 289540
E--> 2605871 aguugcaga uggccac Same 17
2608142 <210> SEQ ID NO: 289793
2608143 <211> LENGTH: 22
2608144 <212> TYPE: RNA
2608145 <213> ORGANISM: Human adenovirus E
2608147 <400> SEQUENCE: 289793
E--> 2608148 cugaauuag cauucuaag aa Same 22
2608601 <210> SEQ ID NO: 289844
2608602 <211> LENGTH: 18
2608603 <212> TYPE: RNA
2608604 <213> ORGANISM: Human adenovirus E
2608606 <400> SEQUENCE: 289844
E--> 2608607 gaguugcag auggccac Same 18
2608745 <210> SEQ ID NO: 289860
2608746 <211> LENGTH: 22
2608747 <212> TYPE: RNA
2608748 <213> ORGANISM: Human adenovirus E
2608750 <400> SEQUENCE: 289860
E--> 2608751 gauggugcan gaguguaaca ug Same 22
2609222 <210> SEQ ID NO: 289913
2609223 <211> LENGTH: 21
2609224 <212> TYPE: RNA
2609225 <213> ORGANISM: Human adenovirus E
2609227 <400> SEQUENCE: 289913
E--> 2609228 gcuccagcuuuccaaccu u Same 21
2610068 <210> SEQ ID NO: 290007
2610069 <211> LENGTH: 14
2610070 <212> TYPE: RNA
2610071 <213> ORGANISM: Human adenovirus E
2610073 <400> SEQUENCE: 290007
E--> 2610074 ugcagagug uaac Same 14
2615792 <210> SEQ ID NO: 290643
2615793 <211> LENGTH: 15
2615794 <212> TYPE: RNA
2615795 <213> ORGANISM: Human adenovirus E
2615797 <400> SEQUENCE: 290643
E--> 2615798 cgcaauaca ccccc Same 15

```



## VARIABLE LOCATION SUMMARY

DATE: 04/05/2004

PATENT APPLICATION: US/10/708,952

TIME: 09:43:44

Input Set : N:\efs\10708952\_efs\SEQ\_LIST.txt

Output Set: N:\CRF4\04052004\J708952.raw

Use of n's or Xaa's (NEW RULES):

Use of n's and/or Xaa's have been detected in the Sequence Listing.

Use of &lt;220&gt; to &lt;223&gt; is MANDATORY if n's or Xaa's are present.

in &lt;220&gt; to &lt;223&gt; section, please explain location of n or Xaa, and which residue n or Xaa represents.

Seq#:111408; N Pos. 8  
Seq#:111419; N Pos. 8  
Seq#:111685; N Pos. 7  
Seq#:111755; N Pos. 11  
Seq#:111789; N Pos. 10  
Seq#:111851; N Pos. 10  
Seq#:112142; N Pos. 5  
Seq#:112305; N Pos. 8  
Seq#:112463; N Pos. 6  
Seq#:113329; N Pos. 7  
Seq#:113436; N Pos. 12  
Seq#:113548; N Pos. 9  
Seq#:289222; N Pos. 11  
Seq#:289540; N Pos. 7  
Seq#:289793; N Pos. 12  
Seq#:289844; N Pos. 8  
Seq#:289860; N Pos. 10  
Seq#:289913; N Pos. 10  
Seq#:290007; N Pos. 5  
Seq#:290643; N Pos. 6  
Seq#:290804; N Pos. 8  
Seq#:290863; N Pos. 9  
Seq#:291145; N Pos. 8  
Seq#:291172; N Pos. 7

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/708,952

DATE: 04/05/2004

TIME: 09:43:44

Input Set : N:\efs\10708952\_efs\SEQ\_LIST.txt

Output Set: N:\CRF4\04052004\J708952.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:1002683 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:111408

L:1002782 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:111419

L:1005176 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:111685

L:1005806 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:111755

L:1006112 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:111789

L:1006670 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:111851

L:1009289 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:112142

L:1010756 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:112305

L:1012178 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:112463

L:1019972 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:113329

L:1020935 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:113436

L:1021943 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:113548

L:2603009 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:289222

L:2605871 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:289540

L:2608148 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:289793

L:2608607 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:289844

L:2608751 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:289860

L:2609228 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:289913

L:2610074 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:290007

L:2615798 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:290643

L:2617247 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:290804

L:2617778 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:290863

L:2620316 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:291145

L:2620559 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:291172